

题目要求：

1. 请继续熟悉 PLY 基本编写规范

仔细阅读 PLY 使用手册

2. 熟悉 yacc\_example 中分析四则运算的程序

3. 编写程序，计算化学分子式中元素的数目，并完成以下测试：

```
atom_count("He") == 1
atom_count("H2") == 2
atom_count("H2SO4") == 7
atom_count("CH3COOH") == 8
atom_count("NaCl") == 2
atom_count("C60H60") == 120
```

参考语法

```
species_list : species_list species
species_list : species
species : SYMBOL
species : SYMBOL COUNT
```

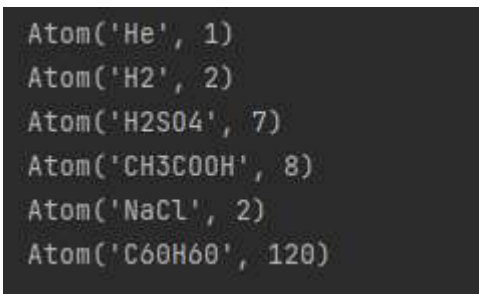
程序说明：

1. 打开 ExpYacc.py 文件，确保和 ExpLex.py 在同一目录下

2. 确保已经安装了 PLY 库

3. 运行 ExpYacc.py 文件

4. 对题目要求中的化学表达式内容进行分析，解析结果如下图：



```
Atom('He', 1)
Atom('H2', 2)
Atom('H2SO4', 7)
Atom('CH3COOH', 8)
Atom('NaCl', 2)
Atom('C60H60', 120)
```

显然解析结果和题目中需要完成的测试结果等价。

5. 对程序 token 定义的解释

根据化学式内容，对出现在解析中的 token 做了相关定义：

```
tokens = (
    'SYMBOL',
    'SYMBOL_COUNT',
)
```

分别表示化学符号和该化学符号对应的分子数。

SYMBOL 和 SYMBOL\_COUNT 的值如下：

```
t_SYMBOL = (
    r"C[laroudsemf]?|Os?|N[eaibdpes]?|S[icernbmq]?|P[drmtboau]?|"
    r"H[eofgas]?|A[lrsgutcm]|B[eraik]?|Dy|E[urs]|F[erm]?|G[aed]|"
    r"I[nr]?|Kr?|L[iaur]|M[gnodt]|R[buhenaf]|T[icebmalh]|"
    r"U|V|W|Xe|Yb?|Z[nr]")

def t_SYMBOL_COUNT(t):
    r'\d+'
    t.value = int(t.value)
    return t
```

## 6. 处理该问题时遇到的关键点——如何累加分子数

其实只要理解了 Yacc 的工作原理即可，关键在于让我们需要的输出结果即是最后总的分子数，于是遇到单个分子的就直接+1，遇到多分子的就+n(n 为 SYMBOL 后面的 SYMBOL\_COUNT)，最后封装一下结果，对于题目要求的那几个化学式进行批量输出就行了。

附图：

```
def p_op1(p):
    'species_list : species_list species'
    p[0] = p[1] + p[2]

def p_op2(p):
    'species_list : species'
    p[0] = p[1]

def p_op3(p):
    'species : SYMBOL'
    p[0] = 1

def p_op4(p):
    'species : SYMBOL SYMBOL_COUNT'
    p[0] = p[2]

def p_error(p):
    print("Syntax error in input!")
```

```
def atom_count(s: str) -> int:
    parser = yacc.yacc()
    result = parser.parse(s)
    return result

if __name__ == '__main__':
    string_list = ['He', 'H2', 'H2SO4', 'CH3COOH', 'NaCl', 'C60H60']
    for s in string_list:
        a1 = Atom(s, atom_count(s))
        print(a1)
    pass
```